Phân Cảnh 3:

Mô hình thuật toán Random Forest là một kỹ thuật học máy khá phổ biến được sử dụng trong cách bài toán hồi quy và phân lớp. Và tại sao nó được gọi là rừng ngẫu nhiên thì một trong những suy nghĩ cơ bản nhất của con người khi nghĩ về rừng là phải có cây mà trong Machines Learning cây ở đây được xác định là một thuật toán là cây quyết định

(nếu trước có nhóm thuyết trình về cây quyết định rồi thì nói :” và như bài diễn giải về mô hình cây quyết định của nhóm xx trước chúng ta sẽ phần nào hiểu được cách nó hoạt động nên chúng tôi sẽ đi sơ qua nó bằng một ví dụ không sâu về nó “ )

(nếu chưa thì nói :” Nếu các bạn ở đây chưa biết cây quyết định là gì thì mô hình sau đây sẽ giúp hiểu khái quát về cách hoạt động của chúng”)

Phân Cảnh 4:

Và đây bên trái của tôi là một mô hình cây quyết đinh cơ bản nhất và bên phải là tập dataset mà tôi dùng để import vào cơ sở dữ liệu. như những gì bạn đang thấy tôi đang cố gắng dự đoán mức lương của nhân viên ở một số bộ phận của ba công ty lớn dựa trên những tính năng nhất định chúng tôi xây một cây quyết định như thế mày, và bây h nhìn chung cây quyết định khá rườm rà và thô kệch. Nếu có thể tôi sẽ biểu diễn nó dưới mô hình rút gọn đơn giản hơn

Phân Cảnh 5 - 6:

Và tôi đã có nó tôi đã tối giảng về mô hình nhị phân, tôi chia nó ra thành 2 mẫu xanh và đỏ và dựa vào đó tôi có thể xây dựng được nhiều cây quyết định từ tập dữ liệu này

Phân Cảnh 7:

Sau đó tôi bắt đầu chia chúng thành một loạt tập dữ liệu ngẫu nhiên và bắt đầu xây dựng hoàn thành được quá trình xây dựng thì nó đã là một khu rừng ngẫu nhiên và tôi cũng sở hữu rất nhiều cây quyết định. Mục đích sử dụng của nó là bạn đưa ra vấn đề bạn muốn dự đoán và tất cả một sẽ đưa ra quyết định, bạn chỉ cần loại bỏ thiểu sô và tham gia đa số thì bạn có thể đưa ra được quyết định của mình. Đây là mấu chốt cơ bản đằng sau mô hình thuật toán RandomForest và dễ hiểu hơn tôi sẽ tiếp cận thực tế trong trường hợp cụ thể sau đó áp dụng RF để tìm ra quyết định

Phân Cảnh 8:

Để mua một cái laptop tôi tiến hành thực hiên một cuộc đến một nhóm bạn của tôi gồm 3 phần tử khủng bố và Vâng bằng cách nào đó họ đã giúp tôi đưa ra được quyết định

Phân Cảnh 9:

những anh chàng thông minh trong ảnh và họ cho tôi một ý kiến khác. Một người trong số họ nói với tôi mua Asus Vì nó được thiết kế để chơi và học, một người khác nói với tôi mua HP vì nó rẻ hơn những chiếc Laptop khác và tôi có thể tiết kiệm được một khoản tiền lớn, tôi có thể dùng số tiền này để mời họ một bữa tối người còn lại thì nói rằng nếu bạn muốn trở thành Kỹ Sư Phần Mềm, tôi nên mua một chiếc Macbook vì nó sinh ra để hữu ích cho bạn. Sau một hồi tranh luận giữa họ, tôi đơn thuần chỉ là một khán giả và tất cả những việc tôi làm đơn giản chỉ là gặp họ trên Google Meet Hay Messenger , thiết lập vấn đề và tham gia vào một cuộc chiến luôn theo số đông

Phân Cảnh 10:

Và Ok không phải lúc nào nó cũng hoạt động trong mọi trường hợp nhưng ít nhất là trường hợp này nó vẫn hoạt động tốt và tôi đã có được quyết định mặc dù tôi không phải là người tạo nên quyết định này tôi chỉ là người thiết lập nên cở sở để tạo nên quyết định này. ở trên tại sao tôi nói không phải lúc nào nó cũng hoạt động trong mọi trường hợp vì không có thuật toán mà mà có độ chính xác ở ngưỡng tuyệt đối 100%

Phân Cảnh 11:

Nếu một thuật toán ở ngưỡng chính xác tuyệt đối thì sẽ không có hiên tượng này xảy hay là hiện tượng này

Phân Cảnh 12:

Về mặt bằng chung một thuật toán tốt chỉ và chỉ khi áp dụng và trường hợp gọi là sở trường của nó khi đó việc dự đoán trung bình có thể mang lại một kết quả tích cực khoảng trên 55% tất nhiên ở đây đang nói về mô hình thuật toán đang làm việc trên một Big Data. Yếu tố con người cũng là một vấn đề kiến thuật toán đưa ra độ chính xác thấp

Phân Cảnh 13:

Đây là một ví dụ điển hình, con người nhờ sỡ hữu bộ nhận thức và bộ sáng tạo không ngừng nên luôn luôn là một biến số khiến các khác thuật toán mới luôn luôn ra đời nhằm đáp ứng biến số thay đổi liên tục mang tên con người.

Và phần diễn giải về mô hình RandomForest tới đây là hết tiếp theo là Logistic Regression